Clayton Lemons
Trevor Young
February 7, 2013

Group Project Proposal

Description

Automatic Emotion Detector (AED) is our proposed idea for using an ANN to learn a small set of emotions for human faces. This emotion set will consist of the following emotions: HAPPY, SAD, ANGRY, SURPRISED, SCARED, and NEUTRAL. AED will use the backpropagation learning algorithm to learn these emotions over a data set of labeled images. Using image pre-processing, AED will be robust enough to detect the emotions in faces of different sizes and orientations.

Data Set

The data set for this task will consist of images containing human faces. The faces will express one of the emotions listed above. We will employ the following two methods for gathering the data set:

1) Using Facebook and email, we will ask friends and family to send pictures of their faces that express each of the emotions in our emotion set. If they are willing, we will ask them to send another set of images expressing the same emotions, but in a slightly different way. This will hopefully allow AED to better learn the variations of each emotion. We will ask each participant to label their own images.

2) We will obtain permission to set up a small booth on campus (possibly in the Wilkinson Center or in the Talmage building) where we can take pictures of volunteers. Like the first method, we will obtain at least one image for each emotion, and we will label the images as we take them. To encourage volunteers, we will reward them with candy or something similar.

Features

To obtain the features of each image, we will apply a chain of pre-processing steps. First, we will use an open source computer vision library called OpenCV to detect, crop, scale, and rotate the faces in the data set images. This will create a consistent data set from which we can extract features. Next, we will convert the images to 8-bit gray scale, simplifying the values of the features. Finally, we will reduce the resolution of the images to limit the number of pixels. Since the normalized intensity value (normalized by 256) of a pixel will make up an input feature, we do not want so many pixels that the training time becomes infeasible.

Feature Example

This example illustrates the pre-processing steps used to obtain the desired input features and shows what an actual feature set would look like. Images A and B are 370×500 pixels in size, and Image C is 37×50 pixels in size. Since the pixel intensities in Image C are used for the input features, there are 1850 features for this example. Only a small subset of the actual features is shown in Table 1.

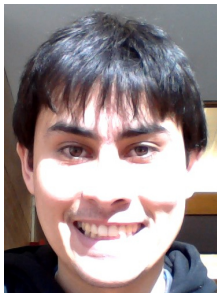

Image A: After cropping

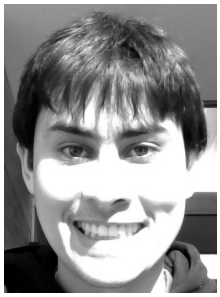

Image B: After grayscale conversion

Image C: After resolution reduction

Feature:	0	1	2	3	4	5	6	7	8	9
Value (Normalized Pixel Intensity):	0.55094	0.55849	0.55849	0.55849	0.55849	0.55849	0.55849	0.55849	0.56226	0.55849

Table 1: The first ten features of Image C. The feature order corresponds to the order of the pixels vectorized in row-major order.

We would both like to be part of this group project.